基于深度学习的四川盆地丘陵区县域耕地遥感识别研究

李　豪[1*]，杜雨秋[2,3]，肖星竹[1]，陈彦羲[1]

（1. 四川农业大学 资源学院，四川成都 611130，中国；2. 中国科学院地理科学与资源研究所，北京 100101，中国；3. 中国科学院大学，北京 101408，中国）

摘　要：［目的/意义］耕地识别是农业遥感的重要应用领域之一，但现阶段深度学习等人工智能方法在中国西南丘陵区遥感识别的研究应用深度仍然不够，识别精度有待提升。［方法］为了快速、精确地获取耕地面积、分布等信息，基于高分辨率的高分六号（GF-6）遥感影像，运用UNet++、DeeplabV3+、UNet与PSPNet等新型深度学习模型对四川省绵阳市三台县耕地信息进行识别，并对各深度学习模型、传统机器学习方法——随机森林法及新型土地覆盖产品SinoLC-1的识别精度进行对比分析，以期深入探索深度学习方法在地物遥感识别领域的应用前景。［结果和讨论］深度学习模型在F_1分数、整体精度（Overall Accuracy, OA）、Kappa系数等精度评价指标的表现上，相比于传统机器学习方法和新型土地覆盖产品均有显著提升，精度提升幅度分别可达20%和50%；其中添加了密集跳跃连接技术的UNet++模型的识别效果最好，其F_1分数、交并比（Intersection over Union, IoU）、平均交并比（Mean Intersection over Union, MIoU）、OA值和Kappa系数值分别为0.92、85.93%、81.93%、90.60%和0.80。应用UNet++模型对2种由仅光谱特征以及光谱+地形特征两种不同特征构建的影像进行耕地提取，光谱+地形特征模型的IoU、OA和Kappa 3个指标比仅光谱特征模型分别提高了0.98%、1.10%和0.01。［结论］深度学习技术在应用于高分辨率遥感影像中的耕地识别方面展现出显著的实用价值，融合光谱和地形特征可以实现信息互补，能进一步改善耕地的识别效果。本研究可为相关部门更好地管理和利用耕地资源、推动农业可持续发展提供技术支撑。

关键词：深度学习；遥感影像；耕地识别；精度评价；丘陵地区

中图分类号：P237　　**文献标志码：**A　　**文章编号：**SA202308002

引用格式：李豪, 杜雨秋, 肖星竹, 陈彦羲. 基于深度学习的四川盆地丘陵区县域耕地遥感识别研究[J]. 智慧农业(中英文), 2024, 6(3): 34-45. DOI：10.12133/j.smartag.SA202308002

LI Hao, DU Yuqiu, XIAO Xingzhu, CHEN Yanxi. Remote Sensing Identification Method of Cultivated Land at Hill County of Sichuan Basin Based on Deep Learning[J]. Smart Agriculture, 2024, 6(3): 34-45. DOI：10.12133/j. smartag. SA202308002 (in Chinese with English abstract)

0　引　言

耕地资源是社会、经济发展的重要基础。虽然中国幅员辽阔，土地资源丰富，但是由于中国地形多为山地丘陵，不利于耕种，耕地资源仅占国土总面积的14.05% [1]。现阶段中国耕地资源形势严峻，耕地面积呈下降之势，耕地质量堪忧、集约利用水平低下、耕地后备资源匮乏等问题日益严重 [2]。因此，为了充分利用和保护耕地，确保国家的粮食安全，实现对生态环境的有效改善，为土地的永续利用打下坚实的基础，实时、精确地获取耕地面积、分布等信息显得尤为重要 [3]。

现阶段获取耕地信息的方法主要有2种：一是传统的实地调查法。该类方法技术成熟，具有精度高、稳定可靠等优点，但仍存在成本较高，受调查方案设计合理性影响较大，数据精度存在显著差异等较多问题，并且通过该方法获得的可用耕地信息十分有限，无法满足时效性和精度的要求 [4]。二是遥感监测方法。遥感技术具有信息获取时效性强且信息量大、探测范围大、受地面条件限制少、途径多样等优点，能有效弥补实地调查法的不足之

收稿日期：2023-07-28

基金项目：国家自然科学基金项目（41501291）；国家级大学生创新训练计划项目（202110626010）；四川天府新区乡村振兴研究院“揭榜挂帅”项目（XZY1-14）

*通信作者：李 豪，博士，研究方向为地理大数据智能分析。E-mail：lihao@sicau.edu.cn

处[5]，目前在耕地信息调查方面已得到广泛应用。

为了减少人力物力的消耗，遥感耕地监测须采用自动化的方法从海量数据集中获取高精度的信息。但传统基于机器学习的遥感信息提取方法往往需要手动设计特征，容易出现特征选择不合理、特征工程不充分等问题，且其通常采用浅层模型，难以捕捉复杂的非线性关系，导致地物的提取效果往往会受到限制[6]。随着计算机算力的不断提升、大数据的普及以及新理论和算法的出现，目前深度学习技术已得到广泛发展。2010年，Minh和Hinton[7]首次将深度学习技术用于获取高分辨率遥感影像的道路信息。2015年，Shelhamer等[8]提出了从根本上改变传统图像语义分割的全卷积网络结构（Fully Convolutional Network, FCN），实现了像素级别的影像分类。此后深度学习技术逐渐被应用于遥感影像的地物识别、变化监测等任务中。但上述深度学习模型仍然存在不足，如没有充分考虑像素与像素之间的关系，缺乏空间一致性；感受野太小，无法获取全局信息等，使得其对图像的分割精度仍受到较大限制。近年来，涌现了如UNet[9]、Deeplab[10-13]、PSPNet[14]等一系列新型深度学习语义分割模型，这些模型均是学者基于深度学习网络进行的创新与改进，也已被应用于耕地等多个地物的遥感识别研究当中[15-17]。

以像元为基础进行遥感地物识别的常规方法主要考虑的是不同地物的影像光谱特性，识别过程中往往没有充分利用地物的其他特征信息。中国西南地区地形起伏大、耕地分布破碎，地表“同谱异物”“同物异谱”现象普遍存在，仅依托影像的光谱特征进行地物识别，会导致提取精度仍受到一定的限制。深度学习语义分割方法能够综合运用不同类型识别特征中蕴含的多层次语义特征，为解决上述问题开辟了新的途径[18]。但目前这些新型方法应用于西南丘陵区遥感地物识别的研究还较少，对识别精度的提升效果还有待验证。

综上所述，研究新型深度学习模型对中国西南丘陵区遥感耕地识别精度的提升具有重要意义，有助于更加精准地获取耕地的数量与空间分布特征等信息。为此，本研究选取四川省中部丘陵区绵阳市三台县为研究对象，基于高分辨率的高分六号遥感影像与区域数字高程模型（Digital Elevation Model, DEM）数据，采用4种新型深度学习模型进行耕地识别研究，验证不同深度学习模型及传统机器学习方法在遥感耕地识别方面的精度，并与一套最新土地覆盖数据产品进行精度对比，探索新型深度学习模型在高分辨率卫星影像丘陵区耕地提取中的可行性，以期为遥感大数据时代人工智能技术的快速发展及其在土地资源监测等系列需求中的成熟应用提供参考依据。

1　研究区及数据来源

1.1 研究区概况

本研究选取四川省绵阳市三台县（30°42'34"～31°26'35"N，104°43'04"～105°18'13"E）为研究区（图1），其位于绵阳市东南部、四川盆地中偏西北部，区域总面积2 659 km^2。该县地处川中丘陵区东北部深丘中窄谷区，长江二级支流郪江流经全境，县域最高点为龙树镇博达岭（海拔672.0 m），最低点位于郪江河谷的短沟口（海拔307.2 m）。研究区属亚热带湿润季风气候，多年平均气温16.7 ℃，无霜期年平均283天，年降水量882～1 134 mm。

三台县具有良好的农业生产基础，县域内耕地分布广泛，耕地面积118.2万亩（788.06 km^2），为四川省典型的丘区农业大县。截至2021年末，三台县常住人口为94.33万人，人均耕地面积1.25亩（833.34 m^2）。三台县地物类型多样，耕地分布破碎分散，土地利用具有典型的丘陵地区特性，有利于检验深度学习模型在遥感地物识别方面的精度。

1.2 数据来源及预处理

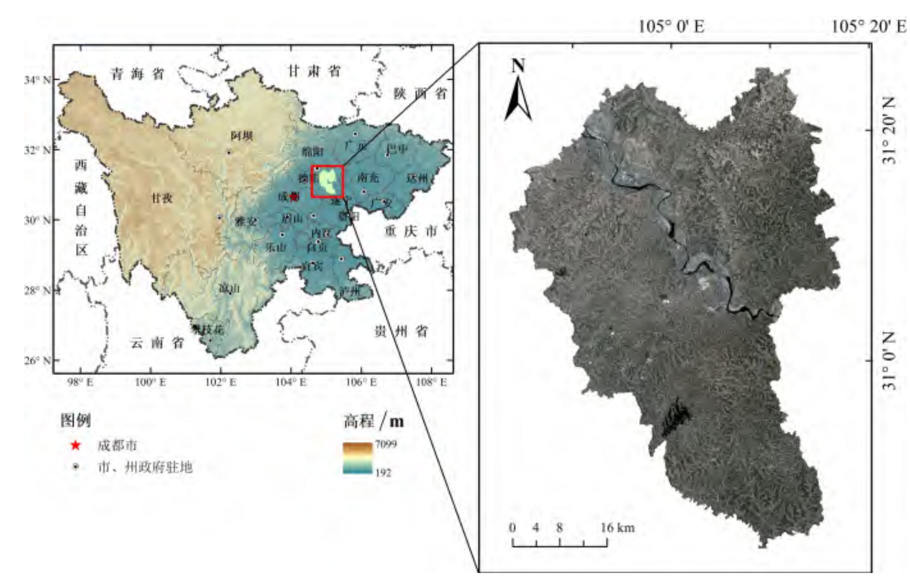
1.2.1　遥感影像数据

本研究使用高分六号（GF-6）卫星影像作为影像数据源，该数据下载自国家航天局对地观测与数据中心（https://www.cheosgrid.org.cn/）。获取了拍摄时间为2020年5月2日的两景影像，云量均少于10%。影像数据概况如表1所示。

遥感数据使用ENVI 5.3软件进行预处理，首先进行辐射定标、FLAASH（Fast line-of-Sight Atmospheric Analysis of Spectral Hypercubes）大气校正等处理；然后基于卫星影像的有理多项式系数（Rational Polynomial Coefficient, RPC）和DEM数据对影像进行正射校正以消除图像几何畸变；再采用Gram-Schmidt融合算法对全色影像和多光谱影像进行融合[19]，融合后多光谱影像分辨率达2 m，有利于更准确提取耕地，为后续的目视解译、样本制作与精度验证等工作奠定基础。

1.2.2　其他数据

研究区的DEM数据来自2020年美国国家航空



注：该图由自然资源部标准地图服务网站下载的审图号为川S【2021】00050号标准地图制作，底图无修改。

图1 四川省绵阳市三台县位置图

Fig.1 Geographical location map of Santai county, Mianyang city, Sichuan province

表1 三台县高分6号卫星影像数据概况

Table1 Overview of GF-6 satellite image data in Santai county

载荷	谱段号	谱段范围/nm	空间分辨率/m
全色相机	Pan	450～900	2
多光谱相机	Blue	450～520	8
	Green	520～600	
	Red	630～690	
	NIR	760～900	

航天局（National Aeronautics and Space Administration, NASA）发布的全新全球30 m分辨率DEM数据（NASA DEM, https://lpdaac.usgs.gov/products/nasadem_hgtv001/）。首先将该数据重采样至与上述卫星影像融合后的空间分辨率（2 m×2 m），然后通过重采样后的DEM数据生成坡度数据，再将DEM、坡度数据与卫星影像的4个光谱特征波段进行波段合成构建共6个波段、用于提取耕地信息的高维特征影像。

本研究还使用了1套土地覆盖数据产品SinoLC-1用于对比评价不同提取方法结果的精度。SinoLC-1数据产品是由武汉大学李卓鸿、中国地质大学（武汉）张洪艳等发布的1 m分辨率国家尺度土地覆盖产品[20]。其基于深度学习框架和包括全球土地覆盖（Global Land Cover, GLC）产品、开放街道地图（Open Street Map, OSM）和谷歌地球图像在内的多源开放访问数据建立而成，总体精度达73.6%。

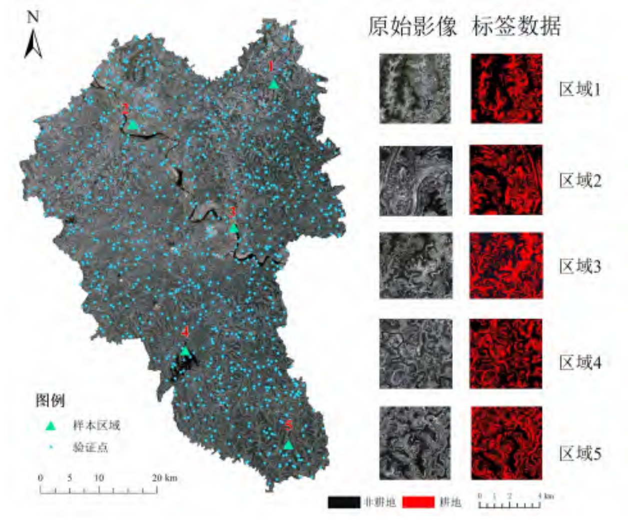
这套数据集的分类系统包含了耕地这一地物类型，可以从中提取耕地信息，用于本研究的精度评价。

2 研究方法

2.1 样本数据集制作

根据研究区地理环境、地势地貌和土地利用等因素，通过人工目视解译获得构建模型所需的地物样本标签数据（Ground Truth, GT），样本尽可能地包含各种地物类型（图2）。

本研究将裁剪重叠率设定为10%，对影像和标



注：该图由自然资源部标准地图服务网站下载的审图号为川S【2021】00050号标准地图制作，底图无修改。

图2 三台县耕地识别模型样本位置及样例

Fig.2 The location and samples of the cropland extraction model in Santai county

签数据同时进行滑动窗口裁剪，得到520组大小为256×256像素的数据集。为了避免训练过程中出现过拟合现象，提高模型的鲁棒性与泛化性，将原始样本数据以5种不同方式（水平翻转、垂直翻转、对角镜像、逆时针旋转90°、顺时针旋转90°）进行数据增强，共得到2 600组样本数据，并以4∶1的比例随机划分为2 080组训练样本和520组验证样本，样例如图2所示。

2.2 深度学习模型

本研究采用Unet、PSPNet、DeeplabV3+以及Unet++共4种深度学习模型开展耕地的遥感识别研究，并结合传统的机器学习方法——随机森林（Random Forest, RF）算法的识别结果，分析不同模型对高分辨率卫星影像的耕地识别精度，为提高耕地利用与调控水平，协调耕地与经济、生态之间的关系提供参考。

1）UNet模型。UNet是Ronneberger[9]于2015年提出的1种对称的U型结构网络，左边编码器由多个卷积层和池化层组成，用于将输入图像降采样并提取特征。右边解码器由多个上采样层和卷积层组成，用于将编码器输出的特征映射恢复到输入图像的分辨率，并且提供逐步细化的分割结果。两者之间设置跳跃连接，以便于解决分割结果中信息丢失的问题。

2）PSPNet模型。Zhao等[14]在2016年提出PSPNet模型，其具有金字塔池化模块（Pyramid Pooling Module, PPM），该模块将特征图分为不同的网格大小，并在每个网格大小上执行池化操作，以捕获相应尺度的上下文信息。最后将不同尺度的池化结果连接在一起，并应用空洞卷积操作以获得最终的上下文特征。

3）DeeplabV3+模型。DeeplabV3+是2018年Chen等[13]在DeeplabV3基础上进行改进后的新型深度学习模型，其将不同空洞率的空洞卷积应用于编码器的多个层中，以获得更大的感受野，构建了由批量归一化（Batch Normalization, BN）层组成的空洞空间金字塔池化层（Atrous Spatial Pyramid Pooling, ASPP），并使用平均池化和最大池化来合并不同的特征信息，通过以级联或并行的方式布局模块。

4）UNet++模型。是Zhou等[21]在2018年基于UNet改进提出了UNet++模型。在UNet的基础上，UNet++模型引入了更多的上采样节点和跳跃连接。该网络中的每个嵌套卷积块通过几个卷积层提取语义信息，通过共享一个编码器将不同深度的UNet有效集成，嵌套卷积块中的卷积层均通过密集的跳跃连接进行连接，使得连接层融合了不同层次的语义信息。这些改进使网络性能得到明显提升。

2.3 机器学习模型

RF模型由Breiman[22]提出，是基于决策树（Classification and Regression Tree, CART）的并行式集成学习算法。通过训练若干CART，然后将各个CART通过一定的组合方式形成RF统计汇总以获得最终结果。RF具有无需做特征选择、不易过度拟合和精度高等优点，是应用广泛的经典机器学习方法。在本研究中，RF模型的输入数据与深度学习模型保持一致，以评价不同模型的耕地识别精度。

2.4 实验环境与实验参数

实验使用Ubuntu20.04.2 LTS操作系统及PyTorch深度学习开发框架，以Python作为开发语言。实验采用的CPU为Intel酷睿i7-12700F，GPU为NVIDIA GeForce RTX 2080 Super。在训练过程中，选用AdamW（Adaptive Moment Estimation with Decoupled Weight Decay）作为优化器，初始学习率设定为0.000 1，模型迭代次数（epoch）设为50，批量大小（batch size）设为4，损失函数采用交叉熵（Cross-Entropy）。

2.5 精度评价

在研究区域内随机生成1 000个矢量验证点构建验证数据集，利用同期0.3 m分辨率的Google地球卫星影像，采用人工目视解译获取验证点所在像元的地物类型，并结合各模型识别结果，通过混淆矩阵（Confusion Matrix）[23]计算整体精度（Overall Accuracy, OA）、交并比（Intersection over Union, IoU）、平均交并比（Mean Intersection over Union, MIoU）、F_1分数（F_1-Score）、Kappa系数（Kappa Coefficient）等5个精度评价指标，对不同模型及数据产品的耕地识别精度进行对比验证。在此基础上，选择精度最高的深度学习模型，分别对包含不同特征的两类影像（识别特征仅包含4个光谱波段（R-G-B-NIR）的影像和在此基础上引入地形信息、共计6个特征波段的影像）进行耕地识别，并计算各精度评价指标以评价在地形复杂的丘陵区不同识别特征对识别精度的影响，具体如

公式（1）～（7）所示。

$$F_1\text{-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{1}$$

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{2}$$

$$\mathrm{MIoU} = \frac{1}{k+1}\sum_{i=0}^{k}\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{3}$$

$$\mathrm{OA} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{4}$$

$$P_0 = \frac{\mathrm{TP} + \mathrm{TN}}{N} \tag{5}$$

$$P_e = \frac{(\mathrm{TP}+\mathrm{FP})\times(\mathrm{TP}+\mathrm{FN})+(\mathrm{TN}+\mathrm{FN})\times(\mathrm{TN}+\mathrm{FP})}{N^2} \tag{6}$$

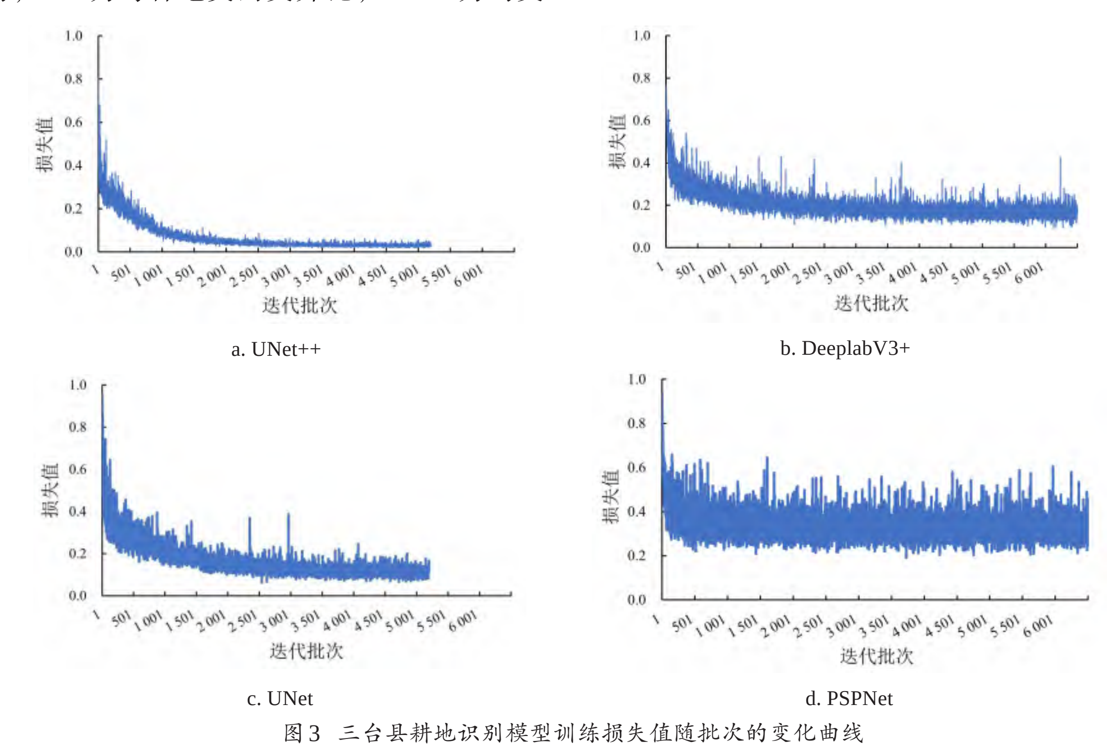
$$Kappa = \frac{P_0 - P_e}{1 - P_e} \tag{7}$$

式中：TP为真实例样本数，个；FP为假实例样本数，个；TN为真反例样本数，个；FN为假反例样本数，个；k为样本类别数，种；N为样本总数，个；*Precision*为精确率，是TP与TP、FP之和的比值，表示预测为真的耕地概率；*Recall*为召回率，是TP与TP、FN之和的比值，表示为被预测为耕地的样本数；F_1-Score是*Precision*和*Recall*的调和平均；IoU为每种地类的交并比；MIoU为均交并比，即指每种地类交并比之平均值；OA表示正确分类的耕地像元数与总的类别个数的比值；P_0为观测一致率；P_e为期望一致率；*Kappa*系数由P_0以及P_e计算得到，是评价判断一致性程度的指标。*Precision*、*Recall*、F_1-Score、OA和*Kappa*系数的取值范围介于0～1之间，数值越大，表明模型的识别精度越高。

3 结果与分析

3.1 模型训练精度对比

如图3所示，UNet++与UNet训练了5 200个batch，DeeplabV3+与PSPNet训练了6 500个batch。在最初的2 000次训练迭代中，4种深度学习模型的验证集损失函数值均呈明显下降趋势，在5 000次batch以后损失值均逐渐形成收敛，最终损失函数值曲线走势平缓，模型表现稳定。这说明实验的学习率设置合理，通过损失函数可实现快速收敛达到最优效果，网络进行了有效学习。4种模型的损失函数值最终分别收敛于0.031 1、0.187 6、0.165 9和0.317 7。其中，UNet++模型经多轮迭代后的准确率水平最高，收敛最快，曲线相对平滑，波动性最小。

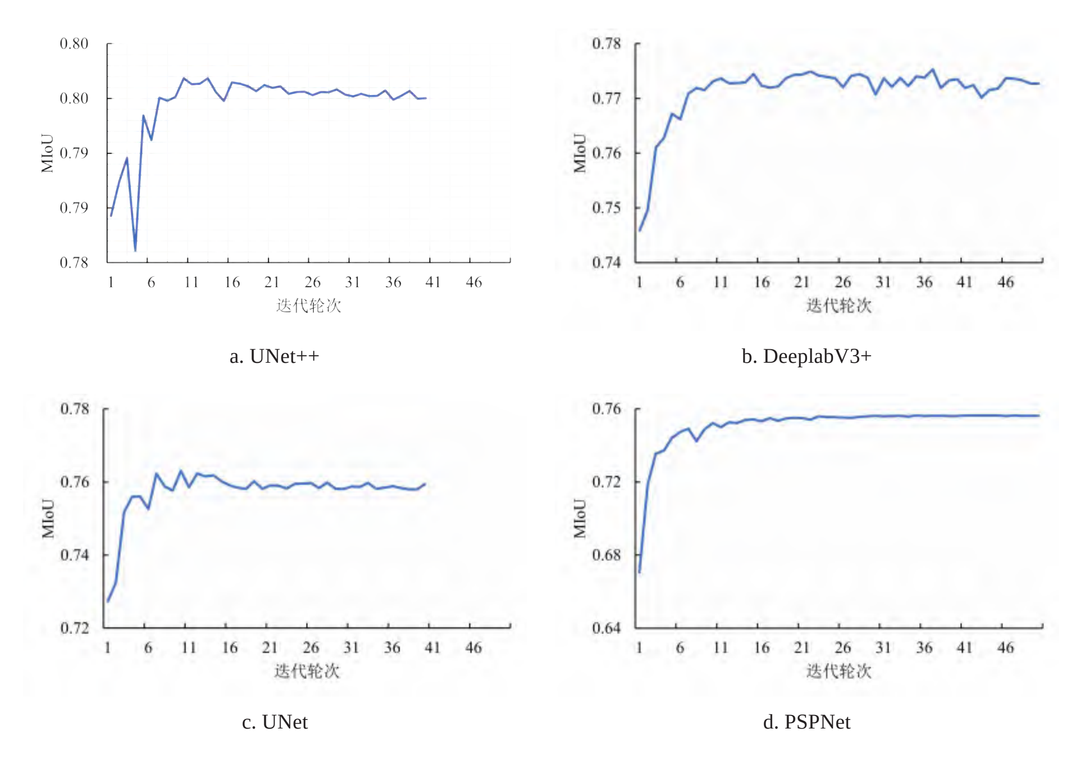


a. UNet++　b. DeeplabV3+　c. UNet　d. PSPNet

图3 三台县耕地识别模型训练损失值随批次的变化曲线

Fig.3 Curve of the fluctuation of train loss of the cropland recognition model in Santai county with batch iterations

在训练过程中验证集MIoU与Kappa系数在迭代50个epoch后，数值不再有明显变化，表明模型学习已经达到收敛，即使再次增加迭代次数，训练精度也不会有较大提升，这表明本实验在迭代50个epoch后可结束训练。由图4和图5的结果可知，UNet++与UNet的收敛速度最快，而PSPNet变化最为平缓，DeeplabV3+最不稳定。4种模型的验证集MIoU最大值分别为0.796 8、0.775 3、0.763 1和0.756 3，Kappa系数最大值分别为0.772 4、0.745 1、0.729 5和0.720 5，两个精度指标的验证结果同样表明UNet++模型在对样本集的耕地识别上具有最好的精度。



a. UNet++　b. DeeplabV3+　c. UNet　d. PSPNet

图4 三台县耕地识别模型MIoU随轮次的变化曲线

Fig.4 Curve of the fluctuation of MIoU of the cropland recognition model in Santai county with epoch iterations

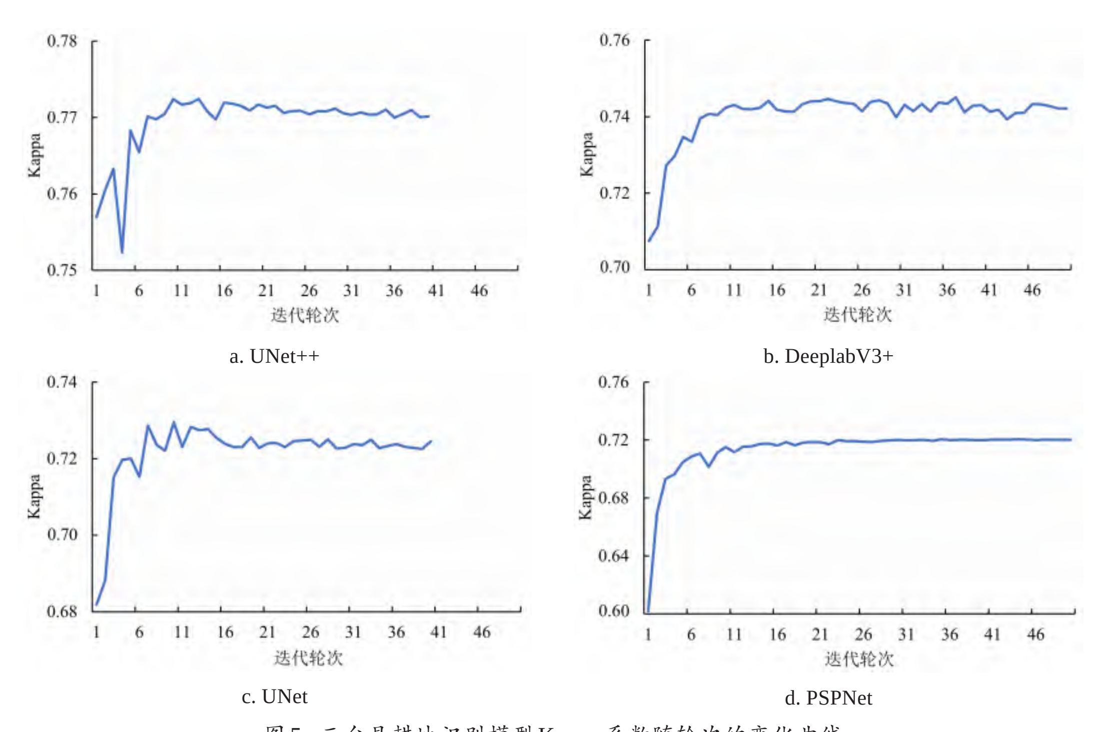
3.2 不同模型识别结果精度对比

3.2.1 精度评价指标对比

基于混淆矩阵的计算结果，计算了验证数据集（1 000个验证点）的MIoU、Kappa系数等5个评价指标（表2），综合各精度评价指标的计算结果可知，各模型/遥感数据产品的耕地识别精度由高到低依次为：UNet++>DeeplabV3+>UNet>PSPNet>RF>SinoLC-1。UNet++、DeeplabV3+模型在F_1-Score、IoU、OA、MIoU和Kappa系数上均优于UNet和PSPNet模型，而UNet在F_1-Score、IoU、MIoU、OA值和Kappa系数上分别比PSPNet模型只提高了0.02、3.12%、3.26%、2.2%、0.05，两者差异不大，说明这两个模型应对当前场景的识别效果处于同一水平。而DeeplabV3+在网络结构设计上借鉴跳步的方式连接低层特征和高层特征，有效实现多尺度特征信息的融合，相比于UNet，从总体识别效果上看，MIoU提高了4.61%，OA值提高了2.7%；在耕地的提取精度上，IoU提高了3.28%。基于UNet改进的UNet++在网络结构上使用了密集的跳跃连接，在解码器子网络上充分抓取了不同层次的信息并进行聚合，从识别效果上看，其F_1-Score、IoU、MIoU、OA值和Kappa系数值分别为0.92、85.93%、81.93%、90.60%和0.80，相比UNet，分别提高了0.03、6.28%、7.35%、4.5%和0.09，在4个深度学习模型中的识别效果最好。此外，可以看出传统的机器学习算法RF法以及土地覆盖数据产品SinoLC-1在耕地提取上效果较差，其精度评价结果均远低于4种深度学习模型，如RF法的IoU、MIoU和OA分别相较UNet++低18.45%、24.6%、15.8%，即相对于这两种提取结果，UNet++模型的精度提升幅度分别可达20%和50%。

3.2.2 识别结果影像对比

三台县卫星影像及各模型/数据产品识别的耕地分布如图6所示。观察研究区的卫星影像可以看



a. UNet++ b. DeeplabV3+

c. UNet d. PSPNet

图5 三台县耕地识别模型Kappa系数随轮次的变化曲线

Fig.5 Curve of the fluctuation of Kappa of the cropland recognition model in Santai county coefficient with epoch iterations

表2 三台县耕地提取结果精度对比

Table2 Comparison of accuracy in extracting cropland in Santai County

模型	MIoU/%	OA/%	F_1-Score	Kappa系数	IoU/%
UNet++	81.93	90.60	0.92	0.80	85.93
DeeplabV3+	79.19	88.80	0.91	0.77	82.93
UNet	74.58	86.10	0.89	0.71	79.65
PSPNet	71.32	83.90	0.87	0.66	76.53
RF	57.33	74.80	0.81	0.45	67.48
SinoLC-1	59.42	75.90	0.81	0.49	67.82

出，三台县内耕地主要集中分布在县域的西部与北部，尤其在郪江两岸阶地上耕地规模更大且更为集中，而南部与西部的耕地分布则较为破碎。

总体而言，各深度学习模型对耕地的识别均表现出良好的整体效果，能够准确地勾勒出耕地分布边界。其中，UNet++模型的耕地总体提取结果最接近真实结果，且在耕地提取完整度、耕地与其他地类的区分及边界提取方面相对于其他模型均具有一定优势。同时，由于遥感影像中地物的复杂性，深度学习模型在进行耕地提取时仍存在一定程度的漏识与错识问题。

与深度学习模型对比，RF法在耕地破碎化程度较高的区域识别表现明显不佳，易将耕地错误地识别为非耕地，大大影响了结果的精度。此外，由于该法属于对影像像元逐一识别的一维分类器，造成识别结果存在一定程度的“椒盐”现象，即明显的孤立点或断点大面积分布[24, 25]。土地覆盖产品的识别结果同样精度不佳，存在大量的错判和漏判情况，制图误差大的区域主要出现在地形破碎、耕地零星散布的区域。

为进一步比较各模型的耕地识别精度差异，选取城镇图7a、水体图7b、林地图7c和连片耕地图7d共4个典型区域，对其耕地识别结果从视觉效果上进行对比分析。这4个区域尽可能代表了研究区域内不同类型的耕地，包括水田及不同种植类型的旱地等，在土地利用、土地覆盖和农作物类型等方面具有显著差异，能够展示各深度学习方法在不同场景下的识别效果。

从土地覆盖产品SinoLC-1的识别结果来看，其采用的卫星影像大多来自亚米级分辨率的Google Earth，随着空间分辨率的提升，其对部分分布破碎的耕地的识别效果较好，该数据产品识别耕地的主要不足在于将大面积的水田混淆为非耕地，并存在将部分水体误判为耕地的问题。

从RF法的识别结果来看，其提取识别效果在总体上优于SinoLC-1，可以在一定程度上检测出不同光谱特征的耕地，如旱地和水田，其主要问题在于对浑浊河流等水体以及部分旱地的识别效果不好，识别结果中将较大范围的含沙河流错误地识别

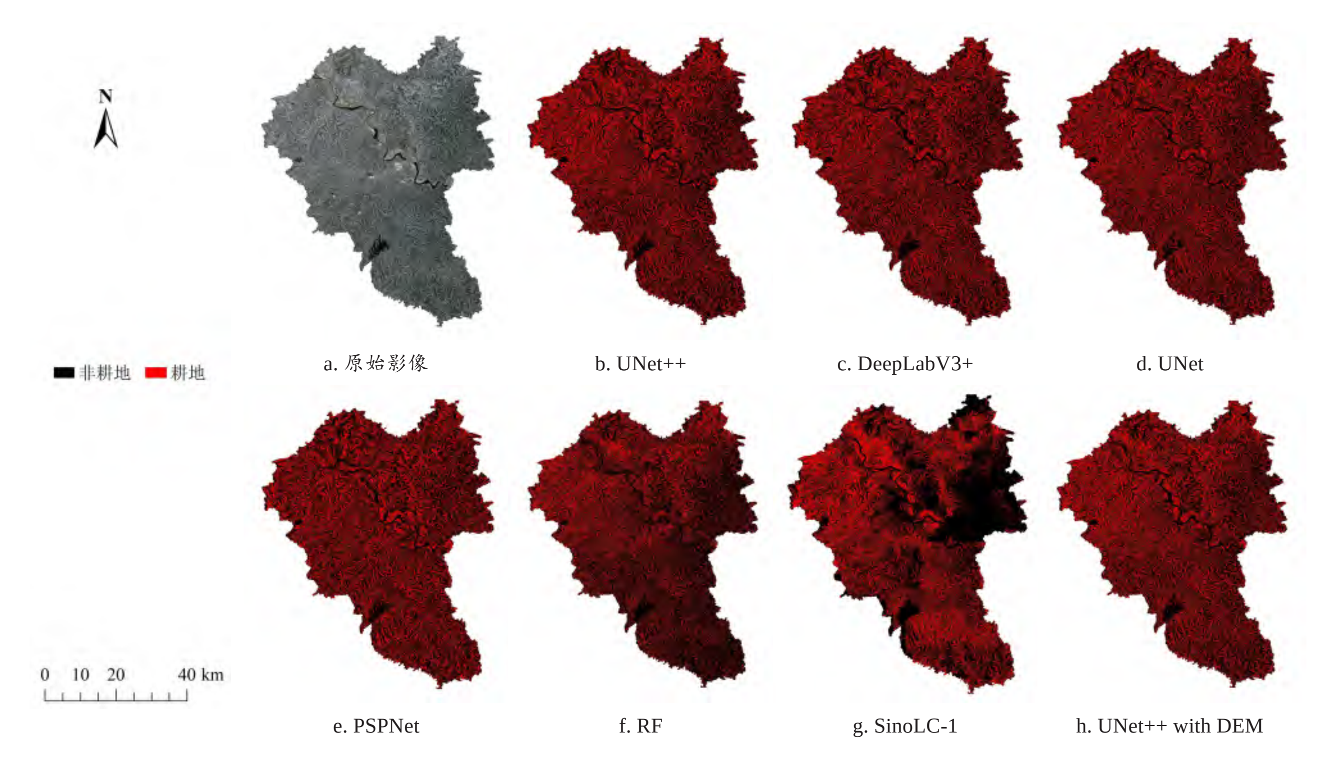


图6　三台县耕地识别结果对比

Fig.6 Comparison of cropland identification results for different models in Santai county

为耕地，总体误差仍较大。

从4种深度学习模型的识别结果来看，对比SinoLC-1和RF的识别结果，各深度学习模型均具有更好的识别精度，特别在水体和分布较破碎的旱地识别方面具有更加精准的优势。其中，PSPNet模型的识别效果逊色于其他深度学习模型，出现了将城镇内的部分道路误识为耕地、对边界清晰的城镇边界提取准确度低、识别结果不连续等较多问题；而DeeplabV3+模型存在对细小地物的分割仍然不够精细等问题。相比UNet模型，改进了网络结构的UNet++模型解决了信息瓶颈问题，有效提高了分割的准确性，其在耕地提取中表现出色，其结果相对其他模型完整，边缘相对其他模型清晰。

3.3 不同特征识别结果精度对比

根据上述分析结果，应用识别精度最高的UNet++模型对2种由不同特征（仅光谱特征和光谱+地形特征）构建的影像进行耕地提取，各评价指标的计算结果如表3所示，可见与仅采用光谱特征的模型相比，引入了DEM、坡度等地形辅助识别特征用于地物识别的深度学习模型可以进一步提高遥感图像的识别精度：例如从IoU、OA和Kappa这3个指标来看，光谱+地形特征模型比光谱特征模型分别提高了0.98%、1.10%和0.01。从上述典型区域的识别结果来看，加入地形特征后的识别效果更优，减少了错判漏判现象，如对零散分布在林地与荒地中的耕地进行了准确的识别，且轮廓较清晰。由此可见，融合光谱和地形特征可以实现信息互补，能进一步改善耕地的识别效果。

4　讨论与结论

4.1 讨论

相对于UNet模型，UNet++模型通过重新设计跳跃连接，在解码器子网络上聚合不同尺度的语义特征，从而创造一种高度灵活的特征融合方案[26]，在本研究中获得了相对较好的识别效果。DeeplabV3+模型通过在编码部分引入包含多个空洞卷积扩张率的多尺度ASPP特征提取模块，增强了模型对不同尺寸的同一物体的识别能力，捕获了更多尺度特征，有利于恢复地物的边缘信息[27]，其识别精度同样较高。相比之下，UNet模型复杂度低，网络层构架较浅，无法明确表征更复杂的地物特征，对细节特征的提取也不完全，造成信息丢失[28]；而PSPNet模型虽然能够提取丰富的高级特征，但随着卷积深度的增加，图像的边缘信息可能会变得模糊，导致无法准确捕捉地物的边界信息[29]。由于遥感影像地物具有尺度多样等特点，

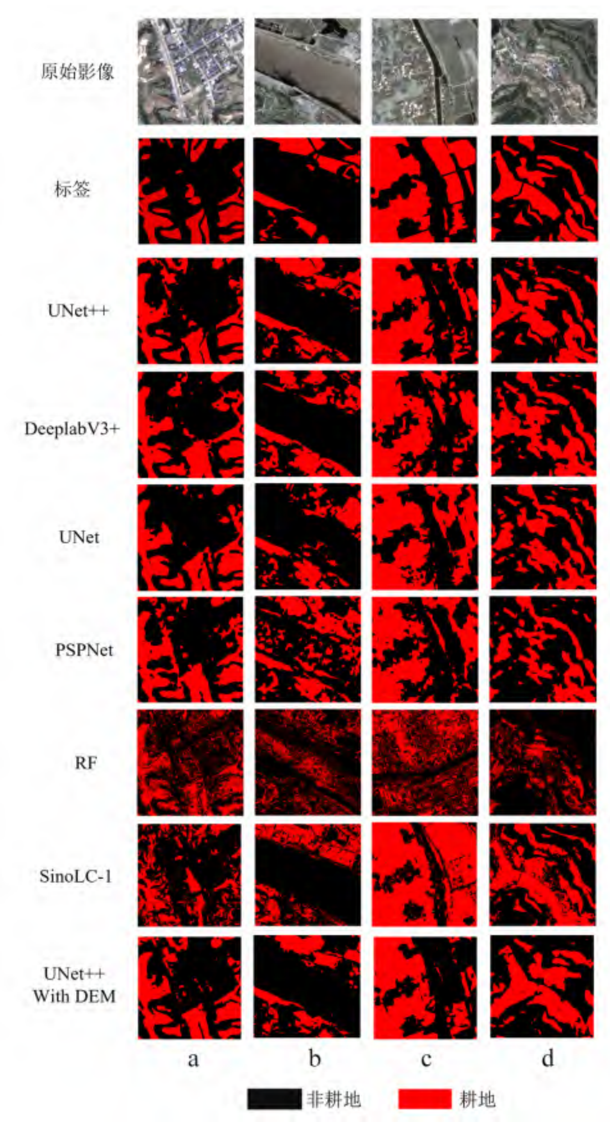


图7 三台县耕地识别局部细节对比

Fig. 7 Comparison of partial details in cultivated land identification in Santai county

表3 UNet++模型对不同特征影像的三台县耕地提取结果精度对比

Table 3 Comparison of accuracy in extracting cropland in Santai county using UNet++ model on different feature images

影像类型	MIoU/%	OA/%	F_1-Score	Kappa系数	IoU/%
光谱特征	81.93	90.60	0.92	0.80	85.93
光谱+地形特征	82.81	91.70	0.94	0.81	86.91

相较于UNet++和DeeplabV3+模型，UNet和PSPNet模型在应对多尺度复杂目标方面的能力相对较差，导致识别精度偏低。

模型的识别效果往往受多种因素的影响，包括数据集的数量与特点、任务的难度和模型的架构等。本研究采用的数据源为分辨率达2 m的高分卫星影像，研究区域选择的是耕地面积小、分布破碎的丘陵地区，从识别结果可以看出，对于此场景中特定的任务和数据集，UNet模型的表现反而优于更晚提出的PSPNet模型。因此，选择合适的模型时，需要根据具体任务和数据集的需求特点进行综合考虑，后续可围绕耕地提取精度随样本量的变化情况进行更加详细的探讨，探究最小样本量等问题。同时，本研究采用的模型均为基础模型，未作改进，后续的研究可向提升模型性能的方向进行探索，如调整编解码器等架构模式、损失函数和扩充特征等。

研究选用的最新土地覆盖产品在本区域存在识别精度偏低，出现较多的错判和漏判等情况，这与前人的研究结果一致[30-32]，说明了此类土地覆盖产品在地形破碎、耕地散布的山地丘陵区的适用性偏低，自行训练模型进行信息提取具有较强的必要性，以更加精准地服务于特定区域的具体应用场景。与传统的机器学习方法（如RF）相比，深度学习模型具有更高的准确性，同样体现了其良好的应用价值。但在研究过程中也发现，仅考虑遥感影像自身的光谱信息等特征来提取耕地并不够全面，而且提取效果很大程度上受到数据质量的影响。此外，对于深度学习方法，训练样本数量是实际应用需求的重要瓶颈之一，直接影响到模型的准确性和应用效率，如何确定最佳的训练数据集规模需要考虑多个因素，这些都应成为今后研究的重点。

4.2 结论

本研究以农业大县四川省绵阳市三台县为研究区，基于高分六号高分辨率卫星影像，采用了Unet、PSPNet、DeeplabV3+以及Unet++4种不同深度学习模型进行了耕地自动提取的实用性和可靠性研究，取得了较为理想的识别效果，主要结果如下。

1）通过OA、Kappa系数等精度指标对耕地提取结果进行了精度评价，发现4种深度学习模型的提取结果均优于机器学习RF方法和新型土地覆盖数据产品SinoLC-1，精度提升幅度分别可达20%和50%。其中，UNet++模型表现最优，精度评价指标OA和Kappa系数分别达到了90.6%和0.80，DeeplabV3+、UNet、PSPNet方法次之。这表明深度学习法在高分辨率遥感影像上快速、准确地获取耕地信息具有一定优势和实用价值。

2）选取城镇、水体、林地和连片耕地4个典

型区域对其耕地识别结果从视觉效果上进行对比分析。PSPNet模型的识别效果逊色于其他深度学习模型，出现了将城镇内的部分道路误识为耕地、对边界清晰的城镇边界提取准确度低、识别结果不连续等较多问题；而DeeplabV3+模型存在对细小地物的分割仍然不够精细等问题。相比UNet模型，改进了网络结构的UNet++模型解决了信息瓶颈问题，有效提高了分割的准确性，其在耕地提取中表现出色，其结果相对其他模型完整，边缘相对其他模型清晰。

3）应用识别精度最高的UNet++模型对2种由不同特征，包括仅光谱特征以及光谱+地形特征，构建的影像进行耕地提取，从IoU、OA和Kappa这3个指标来看，光谱+地形特征模型比仅光谱特征模型分别提高了0.98%、1.10%和0.01，可见，融合光谱和地形特征可以实现信息互补，能进一步改善耕地的识别效果。

根据三台县的耕地提取结果，结合4种深度学习模型的网络结构差异，本研究发现深度学习模型中的UNet++在UNet基础上，引入跳跃连接的机制能够有效提高耕地的提取精度。同时，引入了DEM、坡度等地形特征可以实现信息互补，有助于进一步提高深度学习方法在遥感耕地识别方面的精度。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 刘丹, 巩前文, 杨文杰. 改革开放40年来中国耕地保护政策演变及优化路径[J]. 中国农村经济, 2018(12): 37-51.
LIU D, GONG Q W, YANG W J. The evolution of farmland protection policy and optimization path from 1978 to 2018[J]. Chinese rural economy, 2018(12): 37-51.
[2] 杨翠红, 林康, 高翔, 等. “十四五”时期我国粮食生产的发展态势及风险分析[J]. 中国科学院院刊, 2022, 37(8): 1088-1098.
YANG C H, LIN K, GAO X, et al. Analysis on development and risks of China's food production during 14th five-year plan period[J]. Bulletin of Chinese academy of sciences, 2022, 37(8): 1088-1098.
[3] 于法稳, 代明慧, 林珊. 基于粮食安全底线思维的耕地保护: 现状、困境及对策[J]. 经济纵横, 2022(12): 9-16.
YU F W, DAI M H, LIN S. Cultivated land protection based on bottom line thinking of food security: Current situation, difficulties and countermeasures[J]. Economic review journal, 2022(12): 9-16.
[4] 王佑汉. 基于遥感的四川省撂荒耕地多尺度空间格局及机制研究[D]. 成都: 成都理工大学, 2020.
WANG Y H. Spatial heterogeneity and mechanism of abandoned farmland in different research scales in the Sichuan Province based on remote sense[D]. Chengdu: Chengdu University of Technology, 2020.
[5] 吴培强, 张杰, 马毅, 等. 近20a来我国红树林资源变化遥感监测与分析[J]. 海洋科学进展, 2013, 31(3): 406-414.
WU P Q, ZHANG J, MA Y, et al. Remote sensing monitoring and analysis of the changes of mangrove resources in China in the past 20 years[J]. Advances in marine science, 2013, 31(3): 406-414.
[6] 邴芳飞, 金永涛, 张文豪, 等. 基于机器学习的遥感影像云检测研究进展[J]. 遥感技术与应用, 2023, 38(1): 129-142.
BING F F, JIN Y T, ZHANG W H, et al. Research progress of remote sensing image cloud detection based on machine learning[J]. Remote sensing technology and application, 2023, 38(1): 129-142.
[7] MNIH V, HINTON G E. Learning to detect roads in high-resolution aerial images[C]// Proceedings of the 11th European Conference on Computer Vision: Part VI. New York, USA: ACM, 2010: 210-223.
[8] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(4): 640-651.
[9] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[M]// Lecture notes in computer science. Cham: Springer International Publishing, 2015: 234-241.
[10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computerscience, 2014(4): 357-361.
[11] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 40(4): 834-848.
[12] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// Computer Vision - ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII. New York, USA: ACM, 2018: 833-851.
[13] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. arXiv: 1706.05587, 2017.
[14] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2017: 2881-2890.
[15] LI M M, LONG J, STEIN A, et al. Using a semantic edge-aware multi-task neural network to delineate agricultural parcels from remote sensing images[J]. ISPRS journal of photogrammetry and remote sensing, 2023(200): 24-40.
[16] PERSELLO C, TOLPEKIN V A, BERGADO J R, et al. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping[J]. Remote sensing of environment, 2019, 231: ID 111253.

[17] LIU S J, LIU L C, XU F, et al. A deep learning method for individual arable field (IAF) extraction with cross-domain adversarial capability[J]. Computers and electronics in agriculture, 2022(203): ID 107473.
[18] 刘东杰. 联合波谱和地形特征的深度学习梯田提取方法探讨[D]. 兰州: 兰州大学, 2022.
LIU D J. Study on terraced field extraction with A deep learning method combined with both spectral and topographic features[D]. Lanzhou: Lanzhou University, 2022.
[19] 李国清, 柏永青, 杨轩, 等. 基于深度学习的高分辨率遥感影像土地覆盖自动分类方法[J]. 地球信息科学学报, 2021, 23(9): 1690-1704.
LI G Q, BAI Y Q, YANG X, et al. Automatic deep learning land cover classification methods of high-resolution remotely sensed images[J]. Journal of geo-information science, 2021, 23(9): 1690-1704.
[20] LI Z H, HE W, CHENG M F, et al. SinoLC-1: The first 1 m resolution national-scale land-cover map of China created with a deep learning framework and open-access data[J]. Earth system science data, 2023, 15(11): 4749-4780.
[21] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHSH N, et al. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE transactions on medical imaging, 2020, 39(6): 1856-1867.
[22] BREIMAN L. Random forests[J]. Machine learning, 2001, 45(1): 5-32.
[23] 胡乃勋, 陈涛, 甄娜, 等. 基于卷积神经网络的面向对象露天采场提取[J]. 遥感技术与应用, 2021, 36(2): 265-274.
HU N X, CHEN T, ZHEN N, et al. Object-oriented open pit extraction based on convolutional neural network[J]. Remote sensing technology and application, 2021, 36(2): 265-274.
[24] DURO D C, FRANKLIN S E, DUBÉ M G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery[J]. Remote sensing of environment, 2012, 118(2): 259-272.
[25] 冯文卿, 眭海刚, 涂继辉, 等. 高分辨率遥感影像的随机森林变化检测方法[J]. 测绘学报, 2017, 46(11): 1880-1890.
FENG W Q, SUI H G, TU J H, et al. Change detection method for high resolution remote sensing images using random forest[J]. Acta geodaetica et cartographica Sinica, 2017, 46(11): 1880-1890.
[26] 王超, 王帅, 陈晓, 等. 联合UNet++和多级差分模块的多源光学遥感影像对象级变化检测[J]. 测绘学报, 2023, 52(2): 283-296.
WANG C, WANG S, CHEN X, et al. Object-level change detection of multi-sensor optical remote sensing images combined with UNet++ and multi-level difference module[J]. Acta geodaetica et cartographica Sinica, 2023, 52(2): 283-296.
[27] 黄冬青, 徐伟铭, 许文迪, 等. 基于DeeplabV3+网络的高分遥感影像分类[J]. 激光与光电子学进展, 2023, 60(16): 346-355.
HUANG D Q, XU W M, XU W D, et al. High-resolution remote sensing image classification based on DeeplabV3+Network[J]. Laser & optoelectronics progress, 2023, 60(16): 346-355.
[28] 许玥, 冯梦如, 皮家甜, 等. 基于深度学习模型的遥感图像分割方法[J]. 计算机应用, 2019, 39(10): 2905-2914.
XU Y, FENG M R, PI J T, et al. Remote sensing image segmentation method based on deep learning model[J]. Journal of computer applications, 2019, 39(10): 2905-2914.
[29] 李倩楠, 张杜娟, 潘耀忠, 等. MPSPNet和UNet网络下山东省高分辨耕地遥感提取[J]. 遥感学报, 2023, 27(2): 471-491.
LI Q N, ZHANG D J, PAN Y Z, et al. High-resolution cropland extraction in Shandong province using MPSPNet and UNet network[J]. National remote sensing bulletin, 2023, 27(2): 471-491.
[30] 吴文斌, 杨鹏, 张莉, 等. 四类全球土地覆盖数据在中国区域的精度评价[J]. 农业工程学报, 2009, 25(12): 167-173, 407.
WU W B, YANG P, ZHANG L, et al. Accuracy assessment of four global land cover datasets in China[J]. Transactions of the Chinese society of agricultural engineering, 2009, 25(12): 167-173, 407.
[31] 陈逸聪, 邵华, 李杨. 多源土地覆被产品在长三角地区的一致性分析与精度评价[J]. 农业工程学报, 2021, 37(6): 142-150.
CHEN Y C, SHAO H, LI Y. Consistency analysis and accuracy assessment of multi-source land cover products in the Yangtze River Delta[J]. Transactions of the Chinese society of agricultural engineering, 2021, 37(6): 142-150.
[32] 吴宗洋, 蔡卓雅, 郭英, 等. 黄河流域多源遥感土地覆被数据精度评价与一致性分析[J]. 中国生态农业学报(中英文), 2023, 31(6): 917-927.
WU Z Y, CAI Z Y, GUO Y, et al. Accuracy evaluation and consistency analysis of multi-source remote sensing land cover data in the Yellow River Basin[J]. Chinese journal of eco-agriculture, 2023, 31(6): 917-927.

Remote Sensing Identification Method of Cultivated Land at Hill County of Sichuan Basin Based on Deep Learning

LI Hao[1*], DU Yuqiu[2,3], XIAO Xingzhu[1], CHEN Yanxi[1]

(1. *College of Resources, Sichuan Agricultural University, Chengdu* 611130, *China*; 2. *Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing* 100101, *China*; 3. *University of Chinese Academy of Sciences*, *Beijing* 101408, *China*)

Abstract:

[Objective] To fully utilize and protect farmland and lay a solid foundation for the sustainable use of land, it is particularly important to obtain real-time and precise information regarding farmland area, distribution, and other factors. Leveraging remote sensing technology to obtain farmland data can meet the requirements of large-scale coverage and timeliness. However, the current research and application of deep learning methods in remote sensing for cultivated land identification still requires further improvement in terms of depth and accuracy. The objective of this study is to investigate the potential application of deep learning methods in remote sensing for identifying cultivated land in the hilly areas of Southwest China, to provide insights for enhancing agricultural land utilization and regulation, and for harmonizing the relationship between cultivated land and the economy and ecology.

[Methods] Santai county, Mianyang city, Sichuan province, China (30°42'34"～31°26'35"N, 104°43'04"～105°18'13"E) was selected as the study area. High-resolution imagery from two scenes captured by the Gaofen-6 (GF-6) satellite served as the primary image data source. Additionally, 30-meter resolution DEM data from the United States National Aeronautics and Space Administration (NASA) in 2020 was utilized. A land cover data product, SinoLC-1, was also incorporated for comparative evaluation of the accuracy of various extraction methods' results. Four deep learning models, namely Unet, PSPNet, DeeplabV3+, and Unet++, were utilized for remote sensing land identification research in cultivated areas. The study also involved analyzing the identification accuracy of cultivated land in high-resolution satellite images by combining the results of the random forest (RF) algorithm along with the deep learning models. A validation dataset was constructed by randomly generating 1 000 vector validation points within the research area. Concurrently, Google Earth satellite images with a resolution of 0.3 m were used for manual visual interpretation to determine the land cover type of the pixels where the validation points are located. The identification results of each model were compared using a confusion matrix to compute five accuracy evaluation metrics: Overall accuracy (OA), intersection over union (IoU), mean intersection over union (MIoU), F_1-Score, and Kappa Coefficient to assess the cultivated land identification accuracy of different models and data products.

[Results and Discussions] The deep learning models displayed significant advances in accuracy evaluation metrics, surpassing the performance of traditional machine learning approaches like RF and the latest land cover product, SinoLC-1 Landcover. Among the models assessed, the UNet++ model performed the best, its F_1-Score, IoU, MIoU, OA, and Kappa coefficient values were 0.92, 85.93%, 81.93%, 90.60%, and 0.80, respectively. DeeplabV3+, UNet, and PSPNet methods followed suit. These performance metrics underscored the superior accuracy of the UNet++ model in precisely identifying and segmenting cultivated land, with a remarkable increase in accuracy of nearly 20% than machine learning methods and 50% for land cover products. Four typical areas of town, water body, forest land and contiguous cultivated land were selected to visually compare the results of cultivated land identification results. It could be observed that the deep learning models generally exhibited consistent distribution patterns with the satellite imageries, accurately delineating the boundaries of cultivated land and demonstrating overall satisfactory performance. However, due to the complex features in remote sensing images, the deep learning models still encountered certain challenges of omission and misclassification in extracting cultivated land. Among them, the UNet++ model showed the closest overall extraction results to the ground truth and exhibited advantages in terms of completeness of cultivated land extraction, discrimination between cultivated land and other land classes, and boundary extraction compared to other models. Using the UNet++ model with the highest recognition accuracy, two types of images constructed with different features—solely spectral features and spectral combined with terrain features—were utilized for cultivated land extraction. Based on the three metrics of IoU, OA, and Kappa, the model incorporating both spectral and terrain features showed improvements of 0.98%, 1.10%, and 0.01% compared to the model using only spectral features. This indicated that fusing spectral and terrain features can achieve information complementarity, further enhancing the identification effectiveness of cultivated land.

[Conclusions] This study focuses on the practicality and reliability of automatic cultivated land extraction using four different deep learning models, based on high-resolution satellite imagery from the GF-6 in Santai county in China. Based on the cultivated land extraction results in Santai county and the differences in network structures among the four deep learning models, it was found that the UNet++ model, based on UNet, can effectively improve the accuracy of cultivated land extraction by introducing the mechanism of skip connections. Overall, this study demonstrates the effectiveness and practical value of deep learning methods in obtaining accurate farmland information from high-resolution remote sensing imagery.

Key words: deep learning; remote sensing images; cultivated land identification; accuracy evaluation; hilly region

Foundation item: The National Natural Science Foundation of China (41501291); The Innovation Training Project for Undergraduates of China (202110626010); The Sichuan Tianfu New Area Rural Vitalization Research Institute's 'Revealing the Leaderboard' Project (XZY1-14)

***Corresponding author:** LI Hao, E-mail: lihao@sicau.edu.cn